Advancements in Deep Learning for Tuberculosis Detection: A Review of Modified CNN Architectures and Transfer Learning Approaches for Chest X-ray Analysis

Sonali Nandish Manoli[1], Shaiz Ahmed Shariff[2], Kashyap Manjunath Kaliyur[2], Chaitra Raj Chakra[2], Divith S[2]

[1]Assistant Professor, Department of Computer Science and Engineering (AI&ML),
Vidyavardhaka College of Engineering, Mysore, India
[2]Department of Computer Science and Engineering (AI&ML),
Vidyavardhaka College of Engineering,
Mysore, India

***

Abstract - Tuberculosis remains a significant global health challenge, claiming over 10 million lives annually. Chest radiography continues to be a pivotal tool in the detection and diagnosis of tuberculosis. However, the reliability of human interpretation is hindered by inter-observer variability. To address this, researchers have actively explored deep learning-powered algorithms to enhance the accuracy and efficiency of tuberculosis diagnosis from chest X-rays. This review examines recent research studies that employ deep learning models primarily consisting of modified convolution neural networks with advanced optimization techniques and transfer learning approaches to leverage pretrained networks effectively. Based on these studies, the deep learning models in tuberculosis detection from chest x-rays show promising ranges. Accuracy ranges from 77.14% to 99.6%, sensitivity from 67% to 100%, and specificity from 61% to 100%. F1-scores are between 50% and 99.29%; the area under the curve varies from 0.83 to 0.999, indicating strong model performance across studies. These metrics underscore the high diagnostic potential of deep learning models with modified architectures and transfer learning contributing to improved reliability and generalization in TB classification. This increased reliability will further assist the diagnosis of TB and mitigate the hurdles in treatment caused by the social and economic conditions of the patients.

***Key Words*:** Tuberculosis, Chest X-ray (CXR), Deep Learning, Convolutional Neural Networks (CNN), Transfer Learning, Attention Mechanisms.

1.INTRODUCTION

As per the World Health Organization Global Tuberculosis Report 2024 [1], A global total of 8.2 million people was reported as newly diagnosed with tuberculosis (TB) in 2023, a moderate increase from 7.5 million in 2022 and 7.1 million in 2019 and in sharp contrast to the 5.8 million in 2020 and 6.4 million in 2021. Those newly diagnosed in 2022 and 2023 possibly included a sizable group of people who developed TB in previous years, but whose diagnosis and treatment was delayed by COVID-19 pandemic. TB has possibly returned to being the world's leading cause of death from a single infectious agent. Most of the people who develop TB disease each year are in 30 high TB burden countries, which accounted for 87% of the global total in 2023. Five countries accounted for 56% of the worldwide total: India (26%), Indonesia (10%), China (6.8%), the Philippines (6.8%) and Pakistan (6.3%). One of the barriers to closing diagnostic and treatment gaps is financial costs faced by people with TB and their households. [1] These trends have pushed researchers to develop and make advancements in computer aided diagnosis (CAD) methodologies and we need to first understand the various diagnostic techniques involved. The diagnosis techniques for TB vary significantly in both efficiency and cost. Researchers recognize that advanced imaging methods like computed tomography (CT) and magnetic resonance imaging (MRI) provide high accuracy in diagnosing pulmonary TB due to their three-dimensional imaging capabilities. However, these methods are often costly and challenging to implement in low-resource settings, limiting accessibility for patients in lower-income regions. In contrast, chest X-rays (CXR) offer a more affordable and accessible diagnostic alternative, although they produce two-dimensional images of inherently three-dimensional structures, which can reduce specificity. Given these trade-offs, CXR remains widely used due to its cost-effectiveness and applicability across diverse income groups, making it suitable for large-scale TB datasets [2]. WHO endorses CXR use for mass screening, helping to identify normal cases and reduce the need for costly confirmatory testing. However, high patient volumes and a shortage of skilled radiologists often lead to significant human error in interpreting CXRs, leaving many cases undiagnosed. To address this, ongoing research aims to develop CAD systems for automated TB detection from CXRs, which could be especially beneficial in underdeveloped and developing regions where patient demand exceeds available medical resources. [3] In this review, we examine recent studies (2018–2024) that leverage deep learning algorithms to diagnose TB using CXR images, particularly modified convolution neural networks and transfer learning methods, and to possibly narrow down the methodologies

2. Methodology

An extensive literature review has been made on the use of deep learning techniques to diagnose TB using CXR in the image form. The papers were selected on the basis of their recency (2018-2024) and the journals they were published in were restricted to peer-reviewed journals. A key factor in the selection of the papers was the point of commonality in the methods utilized in the papers belonging to deep learning algorithms that composed of modified convolutional neural networks (CNNs) with advanced optimization techniques and transfer learning approaches to leverage pretrained networks effectively, this was done to find the best performing algorithms in these two specific domains.

2.1 Transfer Learning

In this study, O. Yadav et al [4] have used a methodology that follows a coarse-to-fine approach, leveraging transfer learning and progressive specialization for effective lung classification in CXR images. First, a deep residual network (ResNet-50) addresses the vanishing gradient problem, allowing data to be passed across layers to optimize a residual mapping. In the transfer learning stage, the researchers proposed beginning with lower-resolution images and gradually progressing to higher-resolution ones. Discriminative Fine Tuning is employed here, where the initial layers retain their general feature-detection abilities (e.g., lines and edges) while the middle and final layers progressively learn dataset-specific structures. This is achieved by assigning low learning rates to early layers and higher rates to later layers, focusing learning on complex patterns in the images. For the coarse model, they trained only the fully connected layers (FCL) without adjusting convolution layers to maintain model generality. In the fine model stage, they then applied this pretrained model to the China dataset for further refinement, increasing the training-to-validation split due to its smaller size. This ensured it could generalize well for lung detection across different datasets, resulting in improved accuracy and specificity without over fitting to any single dataset. Datasets used in this study can be summarized in the following divisions, Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, and has a size of 138 images, consisting of TB positive and normal images i.e. Dataset 1(normal=80, TB=58). Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People’s Hospital, China, and has a size of 662 images, consisting of TB-positive and normal images i.e. Dataset 2 (normal=326, TB=336), NIH Dataset: Contains more than 100,000 medium resolution images of 14 different lung diseases. Out of which, 7 are strong symptoms of Tuberculosis, to be specific, Consolidation, Fibrosis, Infiltration, Mass, Nodule, Pleural Thickening and Pneumonia. Note: The authors employed dataset separation and split the dataset into 3 parts. The NIH dataset was used on a pretrained Resnet50 model to fine-tune it on a generalized dataset. For further medium specialization, the Montgomery dataset was used and, for finer fine-tuning the Shenzhen dataset was used. This paper’s methodology of training the model generally, and then specifically has yielded great results. The coarser-to-finer method was further enhanced by optimization techniques like data augmentation, discriminative fine tuning, restarts and, cycle multiplier which gave the following results: Accuracy measured with raw images: 92.8% and accuracy measured with augmented images: 94.89%.

In another study, Chen CF et al [5] have used a methodology using the Google Teachable Machine (GoogleTM) to detect TB in CXRs. GoogleTM, which employs transfer learning, allows for training with minimal coding and computational requirements. The research involved three models trained on two distinct datasets: Model 1 trained to differentiate TB from normal CXRs, Model 2 trained to distinguish abnormalities from normal CXRs, Model 3 combined the outputs of Models 1 and 2 to provide an averaged predictive value, trained to detect both TB and other abnormalities effectively. The training process involved at least 10 iterations per model, selecting the version with the highest accuracy for final use. The models underwent both internal and external validation. The performance of the models was also compared to that of five experienced pulmonologists on a subset of the validation data. The authors have specifically curated diverse datasets for training and validation where they used a single integrated dataset and then randomly split it into a defined ratio for training and validation sets to perform internal validation and also used external validation datasets to perform external validation. Training Dataset 1: Shenzhen dataset, Montgomery dataset, Kaggle’s RSNA Pneumonia Detection Challenge i.e. TB=348 and normal=3806, Training Dataset 2: ChestX-ray14 dataset i.e. Abnormal=1150 and normal=627, Validation Dataset 1: Collected retrospectively from VGHKS (hospital) i.e. TB=83, Abnormal=83 and, normal=84, Validation Dataset 2: a subset of Validation Dataset 1 i.e. TB 50, normal=33 and, Abnormal=22 This study produced the results and also subjected them to both external and internal validation, where internal validation results were produced using the validation subsets of the training datasets and the external validation results were produced using the validation datasets. Internal Validation: Showed high accuracy for all the models: Model 1 showed sensitivity: 96%, specificity: 98%, and accuracy: 97%. Model 2 showed sensitivity: 92%, and specificity: 92%. Model 3 performed better than Models 1 and 2, achieving the highest accuracy across metrics. External Validation: For TB vs. normal cases, Model 3 achieved Area Under the Curve (AUC) of 0.951 and 0.975 on Validation Datasets 1 and 2, respectively, indicating high accuracy. In the TB vs. normal and other abnormal cases scenario, Model 3’s AUC was slightly lower (0.758 and 0.828) due to the added complexity of non-TB abnormalities, which can mimic TB. Physician Comparison: Model 3’s performance was close to the pulmonologists’ accuracy, with AUC values for Model 3 at 0.828 compared to the pulmonologists’ range of 0.843–0.888. When combined with physician input, the accuracy improved further, demonstrating potential benefits from AI-assisted diagnostics.

In another paper, T. Rahman et al [6] present an approach to detecting TB from CXRs using deep learning, segmentation, and visualization techniques. They tested nine pretrained CNN models, including but not limited to ResNet, DenseNet, and ChexNet, applying transfer learning for classification. They also implemented two U-Net models for lung segmentation, aiming to improve classification accuracy by focusing the models on lung regions alone. The models were evaluated using both segmented and non-segmented images. Visualization with Score-Weighted Visual Explanations for Convolutional Neural Networks (Score-CAM) confirmed that model decisions focused on lung regions, enhancing interpretability and reliability in TB classification. This approach demonstrates the potential for reliable, automated TB screening in resource limited settings. Datasets used in this study can be summarized in the following divisions, Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, and has a size of 138 images, consisting of TB-positive and normal images i.e. Dataset 1 (normal=80, TB=58), Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People’s Hospital, China, and has a size of 662 images, consisting of TB-positive and normal images i.e. Dataset 2 (normal=326, TB=336), Belarus Dataset: The dataset contains 306 images and all images are TB-positive i.e. Dataset 3 (TB=306), NIAID TB Dataset: The dataset contains 2,800 images and all images are TB-positive i.e. Dataset 4 (TB=2800), RSNA CXR Dataset: The dataset contains 3500 images and all images are normal CXR images i.e. Dataset 5 (normal=3500) The study demonstrated that DenseNet201, when applied to segmented lung images, achieved the highest accuracy for TB detection, with a classification accuracy: 98.6%, precision: 98.57%, sensitivity: 98.56%, and specificity: 98.54%. Comparatively, CheXNet performed best on non-segmented images with a 96.47% accuracy. Score-CAM visualizations confirmed that models trained on segmented images focused primarily on lung regions, supporting more accurate and interpretable TB classification. These findings highlight the effectiveness of lung segmentation in enhancing model performance and reliability.

In another study by Nijiati et al [7] presents a deep learning based system for detecting pulmonary TB in CXRs, addressing challenges such as variability in radiologist interpretation and limited healthcare resources in high TB-incidence areas. The researchers trained three deep convolutional neural networks (DCNNs)—ResNet, VGG, and AlexNet—to classify images as TB-positive or non-TB. To focus on lung regions specifically affected by TB, a U-Net model was used to segment lung areas before classification. Dataset used in this study can be summarized in the following division, Combined Dataset: The dataset was collected from two hospitals in Kashgar, Xinjiang, China, and has a size of 9,628 images, consisting of TB positive and Non-TB images i.e. Dataset 1 (Non-TB=4,628, TB=5,000) The study found that the ResNet-based AI model achieved the highest diagnostic accuracy for TB detection on chest X-rays, with 96.73% accuracy, 95.50% sensitivity, and 98.05% specificity. It outperformed the VGG and AlexNet models, each achieving AUC scores above 0.99, and successfully identified TB-affected regions through heatmaps that aligned with clinician-diagnosed areas, demonstrating high diagnostic reliability and interpretability for clinical use.

2.2 Modified Convolutional Neural Networks

In another paper, V. T. Q. Huy et al [8] introduce a novel deep learning model called CBAMWDNet which was designed to improve TB detection in CXR images. The CBAMWDNet model combines the Convolutional Block Attention Module (CBAM) with a Wide DenseNet architecture, where CBAM enhances the network's focus on the regions of interest in the image, effectively capturing key features for accurate TB classification. CBAM consists of channel and spatial attention mechanisms, which enable the model to prioritize important information across the image channels and spatial dimensions. Additionally, CBAMWDNet's Wide DenseNet structure provides multiple dense connections between layers, allowing the model to learn more complex features by using a higher number of filters and efficient feature propagation. Datasets used in this study can be summarized in the following divisions, Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, i.e. Dataset 1 (normal=80, TB=58), Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People's Hospital, China, i.e. Dataset 2 (normal=326, TB=336), Tuberculosis (TB) Chest X-ray Dataset: The dataset was Compiled by Qatar University, University of Dhaka, and collaborators in Malaysia, and has a size of 4,200 images, consisting of TB-positive and normal images i.e. Dataset 3 (normal=3,500, TB=700) The CBAMWDNet model demonstrated strong performance in TB detection, giving accuracy: 98.8%, sensitivity: 94.28%, precision: 98.5%, specificity: 95.7%, and F1 score: 96.35% on the TB dataset. The model showed consistent accuracy and generalization across different datasets, performing better than many state-of-the-art models in key classification metrics. It confirmed its effectiveness for TB diagnosis in chest X-rays.

In another study, V. Sharma et al [9] aim to develop a segmentation-based approach that concentrates only on relevant lung areas, enhancing accuracy and making the model more robust and practical for real-world applications. The study uses deep learning to detect TB in CXR images by segmenting and classifying lung regions. The method first trains a UNet segmentation model on datasets from Montgomery County and Shenzhen Hospital, achieving over 96% accuracy in lung segmentation. The approach demonstrates high diagnostic precision, suggesting its value in low-resource settings as a supportive tool for clinicians. Datasets used in this study can be summarized in the following divisions, Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, and has a size of 138 images, consisting of TB-positive and normal images i.e. Dataset 1 (normal=80, TB=58), Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People's Hospital, China, and has a size of 662 images, consisting of TB-positive and normal images i.e. Dataset 2 (normal=326, TB=336), National Institute of Allergy and Infectious Diseases (NIAID) TB Portal Program Dataset: The dataset was available via Kaggle and created by NIAID, and has a size of 1,400 images, consisting of TB-positive and normal images i.e. Dataset 3 (normal=700, TB=700) The study achieved 96.35% accuracy in lung segmentation using the UNet model, with a Jaccard index of 90.38%, Dice coefficient of 94.88%, and AUC of 0.99. For classification, the Xception model reached 99.29% accuracy, with precision, recall, and F1-score all at 99.29% and an AUC of 0.999. The Gradient-Weighted Class Activation Mapping (Grad-CAM) visualizations showed consistent lesion patterns in TB-positive cases, demonstrating the model's precision and reliability in detecting TB in chest X-rays.

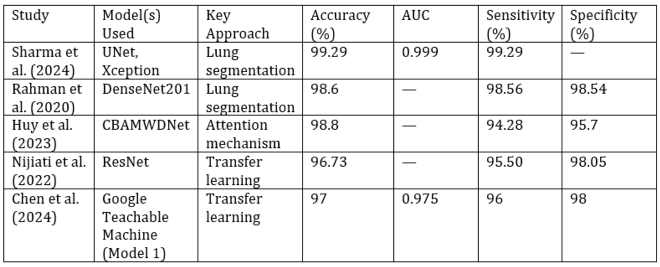
2.3 Modified Convolutional Neural Networks with Transfer Learning

In another study, Meraj et al [10] The researchers implemented four individual CNN models, specifically, VGG-16, VGG-19, ResNet50, and Google Net. These standard CNN architectures were modified uniformly, utilizing pretrained weights. A Global Average Pooling (GAP) layer was added for dimensionality reduction, followed by a dropout layer. The models were evaluated using accuracy and area under the curve (AUC), measuring classification competence through sensitivity and specificity. Datasets used in this study can be summarized in the following divisions, Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, and has a size of 138 images, consisting of TB-positive and normal images i.e. Dataset 1 (normal=80, TB=58), Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People's Hospital, China, and has a size of 662 images, consisting of TB-positive and normal images i.e. Dataset 2 (normal=326,TB=336). The study computed accuracy and AUC for all four models across two datasets. Comparisons were focused on VGG-16, VGG-19, and ResNet50 for consistency. The highest accuracies on dataset 2 were achieved by the VGG-16 and VGG-19 models, outperforming other models evaluated in similar studies. However, the ResNet50 and Google Net models achieved the highest accuracies in previous works on Dataset 2. The metrics of the model are as follows: Accuracy on Dataset 2: VGG-16: 86.74%, VGG19: 84.33%, ResNet50: 81.92%, Google net: 80.72%. AUC on Dataset 2: VGG-16: 92.0%, VGG-19: 91.0%, ResNet50: 91.0%, Google net: 88.0%. Accuracy on Dataset 1: VGG-16: 77.14%, VGG-19: 77.14%, ResNet50: 71.42%, Google net: 71.42%. AUC on Dataset 1: VGG-16: 92.0%, VGG-19: 91.0%, ResNet50: 91.0%, Google net: 88.0%.

P. Rajpurkar et al [11] Using a novel methodology named CheXaid, performed a comparative accuracy study by comparing physicians with and without the assistance of the model in the task of diagnosing TB for HIV positive patients. The model was pretrained on a large dataset (CheXpert), which included 224,316 CXRs of 65,240 patients, with labels for 14 radiological observations not limited to TB, before fine tuning on the training set. They integrated clinical information such as age, oxygen saturation, hemoglobin, CD4 count, white blood cell count, temperature, current antiretroviral therapy status, and the patient's prior history of TB along with CXRs to Improve diagnostic accuracy. Datasets used in this study can be summarized in the following divisions, the first dataset was collected from two secondary level hospitals serving communities with a high burden of HIV and TB in Cape Town, South Africa: GF Jooste Hospital (November 2011–February 2013), and Khayelitsha Hospital (March 2013–October 2014). i.e. Dataset 1 Jooste/Khayelitsha (n=305), The next dataset was collected as part of a diagnostic study of HIV-positive patients with at least one TB symptom admitted to the emergency center of Khayelitsha Hospital from 2016 to 2017 i.e. Dataset 2 Khayelitsha (n = 378). Algorithmic assistance significantly improved physicians' diagnostic accuracy from 0.60 (95% CI 0.57, 0.63) without assistance to 0.65 (95% CI 0.60, 0.70) with assistance, showing a statistically significant increase (chi-square = 9.64, p = 0.002). Sensitivity and specificity both improved with assistance, from 0.70 to 0.73 and 0.52 to 0.61, respectively. However, the stand-alone algorithm outperformed the assisted physicians, achieving a mean accuracy of 0.79 (95% CI 0.77, 0.82) compared to 0.65 (95% CI 0.60, 0.70) for the assisted physicians (chi-square = 66.6, p ¡ 0.0001). The algorithm also demonstrated high specificity (0.87) and sensitivity (0.67) on the same cases viewed with physician assistance. Different training strategies affected the algorithm's performance: on the full test set, it achieved 0.78 accuracy and an AUC of 0.83, while training without clinical covariates or extensive pretraining reduced its accuracy and AUC scores.

A. T. Sahlol et al [12] built a hybrid method that begins with MobileNet, a CNN known for being both powerful and lightweight—perfect for scenarios where resources and computing power might be limited. By starting with a model already trained on ImageNet, a massive dataset with over a million images from various categories, they leveraged the advantages of transfer learning. The researchers used two key datasets for this study, namely, the Shenzhen dataset, which includes 662 CXRs collected from Shenzhen Hospital in China during 2012 i.e. Dataset 1 (Normal=326, TB=336), and a second, larger dataset of 5,232 pediatric chest X-rays, collected and labeled by medical professionals i.e. Dataset 2 (Normal=1,349, TB=3,883). However, dealing with MobileNet's initial extraction of around 50,000 features from these images was a challenge. Processing such a large number of features could slow down the model and increase the risk of overfitting. To address this, the authors implemented the Artificial Ecosystem-based Optimization (AEO) algorithm, which mimics natural processes like how plants grow, animals consume resources, and decomposers recycle nutrients. AEO helped the model intelligently narrow down the feature set, selecting only the most relevant features and discarding the rest. This optimization dramatically reduced the feature set to just 25 features for the Shenzhen dataset and 19 for the second dataset, making the model not only faster but also more generalizable and effective. By taking this approach, the authors struck a balance between achieving high diagnostic accuracy and maintaining efficiency. For dataset 1, the model achieved an impressive accuracy of 90.2%, with a sensitivity of 91.4% and specificity of 90.1%. On dataset 2, the model performed even better, achieving 94.1% accuracy, 87.2% sensitivity, and 97% specificity. This clever combination of deep learning and evolutionary algorithms showcased how medical image analysis could be enhanced, making it both more accurate and practical for real-world healthcare settings.

In another study, by Chirath et al [13] the methodology follows a structured pipeline designed to classify TB in CXR images. First, the dataset was prepared by collecting data of CXRs from standard and publicly available datasets, then, synthetic TB-positive CXR images were generated using a deep generative adversarial network (GAN). The best synthetic images were then selected based on subjective quality assessments by radiologists. These selected images underwent preprocessing to reduce unwanted distortions and enhance relevant features. The lung fields were then segmented using a UNET model using the lung segmentation masks. Next, a DCNN was optimized using a genetic algorithm to finetune its hyperparameters for better classification performance. The final classification was achieved through an ensemble of pretrained VGG16 and InceptionV3 models, which worked together to improve binary classification accuracy for detecting tuberculosis. The model's performance was assessed using metrics such as sensitivity, specificity, and Youden's index. To ensure the model focused on relevant image regions, Class Activation Map (CAM) visualizations were employed to provide insights into the model's decision-making processes. Datasets used in this study can be summarized in the following divisions, Shenzhen Hospital Dataset: The dataset was gathered in Shenzhen No.3 People's Hospital, China, i.e. Dataset 1 (normal=327, TB=335), Montgomery County Dataset: The dataset was gathered in Montgomery County, USA, i.e. Dataset 2 (normal=80, TB=58), Generated Dataset: The authors of this paper have leveraged Generative Adversarial Networks (GAN) to generate 1000 CXRs of TB positive. i.e. Dataset 3 (TB=1000), MIMIC Dataset: The authors further collected 970 CXRs of healthy patients to balance out the dataset i.e. Dataset 4 (normal=970) Segmentation results were promising as the study used a UNET model to segment lungs in CXR images, achieving a high mean Dice Similarity Coefficient (DSC) of 0.989 on the test set. The study employed an ensemble model combining InceptionV3 and VGG16 DCNNs to classify tuberculosis in chest Xray images. The metrics of the ensemble model were as follows, Accuracy: 97.1% (95% CI: 94.1%–99.4%), Sensitivity: 97.9% (95% CI: 93.7%–100%), Specificity: 96.2% (95% CI: 90.9%–99.3%), Youden's Index: 0.941 (highest among models tested). Compared to previous models in the literature (accuracy between 82.35% and 85.3%), the ensemble provided substantial improvement in accuracy. Class Activation Map (CAM) visualizations were also used to confirm that the model focused on relevant features in its classification decisions.



Study	Model(s) Used	Key Approach	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)
Sharma et al. (2024)	UNet, Xception	Lung segmentation	99.29	0.999	99.29	—
Rahman et al. (2020)	DenseNet201	Lung segmentation	98.6	—	98.56	98.54
Huy et al. (2023)	CBAMWDNet	Attention mechanism	98.8	—	94.28	95.7
Nijiati et al. (2022)	ResNet	Transfer learning	96.73	—	95.50	98.05
Chen et al. (2024)	Google Teachable Machine (Model 1)	Transfer learning	97	0.975	96	98

Table -1: Results

4. RESULTS

In literature, it has been observed that the use of modified CNN architectures and transfer learning has yielded impressive results in TB detection from CXRs, achieving near-radiologist accuracy. As seen in Table-1, Studies utilizing segmentation-focused CNNs, such as Sharma et al. [9] with UNet and Xception, reached 99.29% accuracy and an AUC of 0.999 by concentrating on lung regions, while Rahman et al. [6] achieved 98.6% accuracy with DenseNet201 and lung segmentation. Additionally, models with attention mechanisms, like Huy et al. [8] CBAMWDNet, improved feature extraction and recorded a high accuracy of 98.8% and specificity of 95.7%. Transfer learning approaches also contributed to superior performance; for instance, Nijiati et al. [7] attained 96.73% accuracy with ResNet, and Chen et al. [5] leveraged Google Teachable Machine to reach AUCs up to 0.975. These methods highlight the success of modified CNNs and transfer learning in achieving the highest accuracy and robustness in TB classification. We also observed the potential use of generative artificial intelligence methods in Chirath et al. [13] use of GANs to produce high quality synthetic CXR image data as a way to overcome the scarcity of datasets for this use case.

5. DISCUSSIONS

Our findings from this literature review highlight the potential of modified CNNs and transfer learning for bettering TB detection accuracy in CXRs, particularly in environments where resources are limited. Models using segmentation, such as UNet and DenseNet variants, demonstrate that isolating lung regions can significantly improve diagnostic accuracy by reducing unwanted elements and focusing on regions of interest for TB detection. This aligns with the trend in medical imaging where segmentation-based methodologies have been effective in improving model accuracy. The use of attention mechanisms, particularly seen in CBAMWDNet, further emphasizes the advantage of focusing computational resources on regions of interest. Transfer learning has also been pivotal in these advancements as pretrained models like ResNet and Google Teachable Machine adapt effectively to medical imaging tasks by achieving high accuracies despite having smaller datasets. All of these are not without their limitations, as seen with the decreased specificities when non-TB abnormalities are present. This limitation brings forth the need for further improvement in differentiation between TB and other lung pathologies. As this domain continues to evolve, combing CNNs with transfer learning and attention-based architectures holds promise for accurate, robust, and scalable TB detection.

3. CONCLUSIONS

In conclusion, modified CNNs and transfer learning methodologies have shown great promise in improving TB detection using CXRs, achieving accuracies similar to those of radiologists. Segmentation and attention-based methodologies, along 5 with pretrained models, enhance focus on regions of interest, resulting in improved sensitivity and specificity. While limitations remain, especially in differentiating TB and non-TB lung pathologies, these methodologies offer reliable solutions with great potential for deployment in environments with limited resources. Further development in this area could greatly improve TB diagnosis efficiency and provide substantial support to early diagnosis and treatment of TB.

ACKNOWLEDGEMENT

We thank the Department of CSE(AI&ML), Vidyavardhaka College of Engineering for supporting and enabling us to write this literature review

REFERENCES

1. Geneva: World Health Organization: Global tuberculosis report 2024.
2. Santosh, K., Allu, S., Rajaraman, S., Antani, S.: Advances in Deep Learning for Tuberculosis Screening using Chest X-rays: The Last 5 Years Review. J Med Syst. 46, 82 (2022). https://doi.org/10.1007/s10916-022-01870-8.
3. Hooda, R., Sofat, S., Kaur, S., Mittal, A., Meriaudeau, F.: Deep-learning: A potential method for tuberculosis detection using chest radiography. In: 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). pp. 497–502. IEEE, Kuching (2017). https://doi.org/10.1109/ICSIPA.2017.8120663.
4. Yadav, O., Passi, K., Jain, C.K.: Using Deep Learning to Classify X-ray Images of Potential Tuberculosis Patients. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 2368–2375. IEEE, Madrid, Spain (2018). https://doi.org/10.1109/BIBM.2018.8621525.

5. Chen, C.-F., Hsu, C.-H., Jiang, Y.-C., Lin, W.-R., Hong, W.-C., Chen, I.-Y., Lin, M.-H., Chu, K.-A., Lee, C.-H., Lee, D.L., Chen, P.-F.: A deep learning-based algorithm for pulmonary tuberculosis detection in chest radiography. Sci Rep. 14, 14917 (2024). https://doi.org/10.1038/s41598-024-65703-z.
6. Rahman, T., Khandakar, A., Kadir, M.A., Islam, K.R., Islam, K.F., Mazhar, R., Hamid, T., Islam, M.T., Kashem, S., Mahbub, Z.B., Ayari, M.A., Chowdhury, M.E.H.: Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization. IEEE Access. 8, 191586–191601 (2020). https://doi.org/10.1109/ACCESS.2020.3031384.
7. Nijiati, M., Ma, J., Hu, C., Tuersun, A., Abulizi, A., Kelimu, A., Zhang, D., Li, G., Zou, X.: Artificial Intelligence Assisting the Early Detection of Active Pulmonary Tuberculosis From Chest X-Rays: A Population-Based Study. Front. Mol. Biosci. 9, 874475 (2022). https://doi.org/10.3389/fmolb.2022.874475.
8. Huy, V.T.Q., Lin, C.-M.: An Improved Densenet Deep Neural Network Model for Tuberculosis Detection Using Chest X-Ray Images. IEEE Access. 11, 42839–42849 (2023). https://doi.org/10.1109/ACCESS.2023.3270774.
9. Sharma, V., Nillmani, Gupta, S.K., Shukla, K.K.: Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images. Intelligent Medicine. 4, 104–113 (2024). https://doi.org/10.1016/j.imed.2023.06.001.
10. Meraj, S.S., Yaakob, R., Azman, A., Faculty of Computer Science and Information Technology, Faculty of Computer Science and Information Technology, Faculty of Computer Science and Information Technology, Rum, S.N.M., Faculty of Computer Science and Information Technology, Nazri, A.S.A., Faculty of Computer Science and Information Technology, Zakaria, N.F., Department of Medicine, Medical and Health Science Faculty, Universiti Putra Malaysia, 43400, Serdang, Selangor: Detection of Pulmonary Tuberculosis Manifestation in Chest X-Rays Using Different Convolutional Neural Network (CNN) Models. IJEAT. 9, 2270–2275 (2019). https://doi.org/10.35940/ijeat.A2632.109119.
11. Rajpurkar, P., O'Connell, C., Schechter, A., Asnani, N., Li, J., Kiani, A., Ball, R.L., Mendelson, M., Maartens, G., Van Hoving, D.J., Griesel, R., Ng, A.Y., Boyles, T.H., Lungren, M.P.: CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. npj Digit. Med. 3, 115 (2020). https://doi.org/10.1038/s41746-020-00322-2.
12. Sahlol, A.T., Abd Elaziz, M., Tariq Jamal, A., Damaševičius, R., Farouk Hassan, O.: A Novel Method for Detection of Tuberculosis in Chest Radiographs Using Artificial Ecosystem-Based Optimisation of Deep Neural Network Features. Symmetry. 12, 1146 (2020). https://doi.org/10.3390/sym12071146.
13. Dasanayaka, C., Dissanayake, M.B.: Deep Learning Methods for Screening Pulmonary Tuberculosis Using Chest X-rays. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. 9, 39–49 (2021). https://doi.org/10.1080/21681163.2020.1808532.